Personal Autonomy

New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy

Edited by

JAMES STACEY TAYLOR
Louisiana State University

Contents

v

Introduction

James Stacey Taylor

In recent years, the concept of autonomy has become ubiquitous in moral philosophy. Discussions of the nature of autonomy, its value, and how one should respect it are now commonplace in philosophical debates, ranging from the metaphysics of moral responsibility to the varied concerns of applied philosophy. All of these debates are underpinned by an increasingly flourishing and sophisticated literature that addresses the fundamental question of the nature of personal autonomy.

The concept of autonomy has, of course, been important for moral philosophy for some time, being central to the ethical theories of both Immanuel Kant and such contemporary Kantians as Thomas Hill and Christine Korsgaard.[1] However, recent interest in personal autonomy does not focus on the Kantian conception of autonomy on which a person is autonomous if her will is entirely devoid of all personal interests. Instead, it focuses on a more individualistic conception of this notion, whereby a person is autonomous with respect to her desires, actions, or character to the extent that they originate in some way from her motivational set, broadly construed.

Interest in this individualistic conception of autonomy was stimulated by the publication of a series of papers in the early 1970s, in which Harry Frankfurt, Gerald Dworkin, and Wright Neely independently developed "hierarchical" accounts of personal autonomy.[2] The shared core of these accounts is both simple and elegant: A person is autonomous with respect to a first-order desire that moves her to act (e.g., she wants to smoke, and so she smokes) if she endorses her possession of that first-order desire (e.g., she wants to want to smoke). This approach to analyzing autonomy has much to recommend it. First, it captures an important truth about

persons: They have the capacity to reflect on their desires and to endorse or repudiate them as they see fit. Second, it is an explicitly naturalistic and compatibilist approach to analyzing autonomy. As such, it fits well with the currently dominant compatibilist analyses of moral responsibility, and it seems able to disavow the implausible claim that personal autonomy is incompatible with the truth of metaphysical determinism – a disavowal that is defended by Bernard Berofsky and Alfred Mele in their chapters in this volume.[3] Finally, this approach to analyzing autonomy is content neutral, for it does not require persons to hold any particular values in order for them to be autonomous. This enables it to be readily applicable to many debates within applied ethics where respect for autonomy is of primary concern and where this focus on autonomy is driven by the recognition that some means must be found to adjudicate between competing value claims in a pluralistic society.[4]

Yet despite the many advantages of the hierarchical approach to analyzing autonomy, it suffers from significant theoretical difficulties. In the light of these criticisms, some proponents of the hierarchical approach to analyzing autonomy (such as Stefaan Cuypers and Harry Frankfurt) have developed sophisticated defenses of it.[5] Other writers have developed a "second generation" of neohierarchical theories of autonomy that, while they move beyond the hierarchical approach to analyzing autonomy, acknowledge that the origins of their views lie in the original Frankfurt-Dworkin-Neely theory cluster. Two of the most prominent of these neohierarchical theories of autonomy are those developed by John Christman and Michael Bratman. Christman's historical approach retains the hierarchical analyses' requirement that the attitudes of the person whose effective first-order desire is in question are in some way autonomy conferring. However, rather than holding that this person must in some way endorse the desire in question for her to be autonomous with respect to it, Christman holds that she must not reject the *process* that led her to have this desire.[6] Bratman's analysis of autonomy – the key elements of which he outlines in the chapter "Planning Agency, Autonomous Agency" – combines his influential account of intention and planning agency with certain elements of the hierarchical approach to autonomy.[7] Such neohierarchical approaches to personal autonomy have also been joined by a number of diverse and original approaches to analyzing autonomy that depart from the hierarchical approach altogether. These new approaches to analyzing autonomy include, but are not limited to, the coherentist approach of Laura Waddell Ekstrom,[8] the "helmsman" approach of Thomas May,[9] the doxastic approach of Robert Noggle,[10]

the sociorelational approach of Marina Oshana,[11] and the foundational-ist approach of Keith Lehrer.[12] This debate over the nature of autonomy has led to a significant increase in the philosophical understanding of this concept, and so it is no longer correct that outside of the Kantian tradition autonomy "is a comparatively unanalyzed notion," as John Christman was truthfully able to write in 1988.[13] Moreover, the increasing attention that the concept of autonomy has recently received is not only of interest to autonomy theorists. This is because, as I outline in Section IV, which analysis of personal autonomy turns out to be the most defensible will have direct implications for all debates in moral philosophy in which this concept plays a major role.

These, then, are exciting times for both autonomy theorists and all who draw upon the concept of autonomy. The chapters in this volume, each original to it, represent the state of the art of the current discussion of autonomy and the roles that it plays in discussions of moral responsibility and applied philosophy. The purpose of this Introduction, thus, is to provide the theoretical background against which these chapters were written, by outlining the progress of the debate over the nature and role of autonomy as this has unfolded over the past three decades. As such, it can naturally be divided into four sections. The first will provide the theoretical background to this collection as a whole, through outlining Frankfurt's and Dworkin's hierarchical analyses of autonomy together with the major criticisms that have led to their modification. Despite these modifications, however, I will note that even in their most recent forms these analyses are both still vulnerable to serious theoretical objections.

The second section of this Introduction will outline three of the most prominent recent analyses of autonomy that have been developed to avoid the difficulties that beset the Frankfurt-Dworkin-Neely hierarchical approach: John Christman's historical approach, Michael Bratman's reasons-based view, and Laura Waddell Ekstrom's coherentist analysis. The second section of the Introduction will serve as a supplement to the first, as it provides an introduction to the most recent theoretical literature on autonomy. In so doing, it will serve as a useful backdrop to the discussions in the first part of this collection, "Theoretical Approaches to Personal Autonomy," in which Bratman and Ekstrom outline and develop their respective analyses of autonomy and in which the relationships among autonomy, free will, the "self," and the concept of "identification" are considered.

The third section of this Introduction will outline alleged connections between personal autonomy and moral responsibility. This will provide

the theoretical background to the second part of this collection, "Autonomy, Freedom, and Moral Responsibility." Finally, the last section of this Introduction will indicate the various ways in which the concept of autonomy is invoked within areas of contemporary philosophy apart from discussions of moral responsibility. This section will provide a useful basis from which to approach the final part of this book, "The Expanding Role of Personal Autonomy," which focuses on the role that autonomy plays in political philosophy and in various fields of applied ethics.

I. THE HIERARCHICAL ANALYSES OF AUTONOMY

The core feature shared by Frankfurt's and Dworkin's analyses of autonomy and identification is that these concepts are to be analyzed in terms of hierarchies of desire. (For the sake of clarity, I henceforth take the phrase "is autonomous with respect to her desire *x*" to be synonymous with the phrase "identifies with her desire that *x*.")[14] More specifically, on Frankfurt's original analysis of autonomy a person is autonomous with respect to her first-order desire that moves her to act (her "will") if she *volitionally endorses* that desire. (A "first-order" desire is a desire that a particular state of affairs obtains.) That is to say, a person is autonomous with respect to her effective first-order desire that *x* if she both desired to have the desire that *x* (i.e., she had a *second-order desire* that she have her desire that *x*, where a "second-order" desire is a desire about a first-order desire) *and* she *also* wanted her desire that *x* to move her to act (i.e., she endorsed her desire that *x* with a second-order *volition*).[15] Similarly, on Dworkin's original analysis of autonomy an "autonomous person is one who does *his own* thing," where "the attitude that [the] person takes towards the influences motivating him . . . determines whether or not they are to be considered 'his.'"[16] That is to say, on Dworkin's view a person is autonomous with respect to the desires that motivate him if he endorses his being so moved. In addition to requiring that a person's motivations be "authentic" in this way, Dworkin also required that she enjoy both procedural independence and substantive independence with respect to her motivations. A person possesses procedural independence with respect to her motivations if her desire to be moved to act by them has not been produced by "manipulation, deception, the withholding of relevant information, and so on."[17] A person possesses substantive independence with respect to his motivations if he does not "renounce his independence of thought or action" prior to developing them.[18]

On both Frankfurt's and Dworkin's hierarchical analyses, then, a person's autonomy is impaired if she is moved to act by a desire that she does not volitionally endorse – if she has a second-order desire not to be moved by the first-order desire that is effective in moving her to act. In most cases, this is intuitively plausible. For example, if a person is subject to a constant neurotic compulsion to wash his hands from which he desires to be free, then his autonomy will be impaired if he is moved to act by a first-order desire to wash his hands that this neurosis causes him to have and by which he does not wish to be moved. Similarly, if a person is a "wanton," if he does not care which of his desires moves him to act, then it seems plausible to claim that he is not autonomous (he is not "self-directed"), either because his "self" is not engaged in directing his desires or actions or because he has no coherent "self" to play this role.

Yet despite their plausibility, these early hierarchical analyses of autonomy are subject to three serious objections. The first of these is the Problem of Manipulation.[19] Frankfurt's hierarchical analysis of autonomy is an *ahistorical* (or *structural, punctuate,* or *time slice*) account of autonomy, on which a person is autonomous with respect to his effective first-order desires *irrespective* of their historical origins, provided that he volitionally endorses them. The proponents of the Problem of Manipulation note that a third party (such as a nefarious neurosurgeon or a horrible hypnotist) could inculcate into a person both a certain first-order desire (e.g., the desire to smoke) and a second-order volition concerning this desire so that there is the pertinent sort of hierarchical endorsement. Because this inculcated first-order desire would satisfy Frankfurt's conditions for its possessor to be autonomous with respect to it, Frankfurt is committed to holding that she is autonomous with respect to it – but this ascription of autonomy to her with respect to this desire is suspect.[20]

Of course, Dworkin's analysis of autonomy is not directly subject to the Problem of Manipulation because it is blocked by his requirement that the process by which a person comes to have her desires be one that is procedurally independent – a condition that is clearly unsatisfied when a person's desires are inculcated into her through hypnosis or neurosurgery without her consent. Despite this, one can still use the Problem of Manipulation to develop an *indirect* objection to Dworkin's analysis of autonomy. Thus, although Dworkin's requirement of procedural independence enables him to avoid the Problem of Manipulation, it only does so by fiat, by simply ruling *ex cathedra* that a person is not autonomous with respect to those desires that he has been manipulated into possessing. And this is not enough for his analysis of autonomy to be

theoretically satisfactory. This is because an acceptable analysis of auton-
omy should not merely list the ways in which it is intuitively plausible that
a person will suffer from a lack of autonomy with respect to her effective
first-order desires, but must also provide an account of *why* a person's
autonomy would be thus undermined, so that influences on a person's
behavior that do not seem to undermine her autonomy (e.g., advice) can
be differentiated from those that do (e.g., deception).

Frankfurt's and Dworkin's analyses of autonomy also face the
Regress-cum-Incompleteness Problem.[21] On these analyses, a person is
autonomous with respect to her effective first-order desires if she endorses
them with a second-order desire. Because this is so, the question arises as
to whether this person is autonomous with respect to this second-order
desire and, if she is, why this is so. If she is autonomous with respect to
this second-order desire because it is, in turn, endorsed by a yet higher-
order desire, then a regress threatens, for the question will then arise
as to whether she is autonomous with respect to this *third*-order desire –
and so on. If, however, this person is autonomous with respect to the
second-order desire for a reason *other* than its endorsement by a higher-
order desire, then the hierarchical approach to analyzing autonomy is
incomplete.

Of course, the proponents of the hierarchical approach *could* avoid the
Regress-cum-Incompleteness Problem simply by claiming that although
the person in question is *not* autonomous with respect to her higher-
order endorsing desire, she *is* autonomous with respect to her endorsed
first-order desire, because autonomy is simply constituted by such an en-
dorsement. Yet although Frankfurt and Dworkin could avoid the Regress-
cum-Incompleteness Problem by adopting this line of response, neither
of them does so, no doubt because they recognize that were they to do so
they would encounter the equally troubling *Ab Initio* Problem: How can
a person become autonomous with respect to a desire through a process
with respect to which she was not autonomous? Or, in other words, how
is it that a person's higher-order desires possess any authority over her
lower-order desires?[22] When put in this way, the *Ab Initio* Problem is often
termed the Problem of Authority and in this guise has been neatly en-
capsulated by Gary Watson: "Since second-order volitions are themselves
simply desires, to add them to the context of conflict is just to increase
the number of contenders; it is not to give a special place to any of those
in contention."[23]

Faced with these three difficulties, both Frankfurt and Dworkin mod-
ified their original analyses. Recognizing that his analysis would be

subjected to the Regress-cum-Incompleteness Problem, Frankfurt attempted to eliminate the possibility of such a problematic regress by claiming that a person's *decisive* identification with one of his desires would terminate it.[24] Frankfurt elaborated this decision-based version of his hierarchical analysis of autonomy in "Identification and Wholeheartedness," where he argued that a person is autonomous with respect to his effective first-order desire if he *decisively* endorses it with a second-order volition. Directly responding to the Regress Problem, Frankfurt claimed that if a person endorses his effective first-order desire "without reservation . . . in the belief that no further accurate inquiry would require him to change his mind," it would be pointless for him to continue to assess whether he was autonomous with respect to the first-order desire that was in question.[25] Furthermore, a person's decisive identification with his endorsing second-order volition also seems to circumvent the *Ab Initio* Problem/Problem of Authority, for through this decision the person in question will endow his volition with the authority that it previously lacked.

Unlike Frankfurt, Dworkin did not directly attempt to address criticisms of his analysis of what conditions must be met for a person to be autonomous with respect to her desires and actions. Instead, he clarified that his account was concerned *not* with the *local* conception of what conditions must be met for a person to be autonomous with respect to her actions (or desires), but, instead, with a more *global* conception of autonomy as a "second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes and so forth."[26] Dworkin argued that once it is understood that he was not trying to provide an account of what made a person autonomous with respect to her desires or actions, his conception of autonomy avoids the Regress-cum-Incompleteness Problem. This is because, he claimed, as long as a person enjoyed procedural independence with respect to her reflection upon her desires, there would be "no conceptual necessity for raising the question of whether the values, preferences at the second order would themselves be valued or preferred at a higher level. . . ."[27] Similarly, Dworkin held that his account of autonomy is unaffected by the *Ab Initio* Problem/Problem of Authority. Because on his view persons enjoy autonomy when they engage this capacity for reflection, the exercise of this second-order capacity for endorsement *just is* what is involved in being autonomous.

Yet even if Dworkin's more global approach to analyzing personal autonomy avoids the major problems that were outlined above, this is achieved at considerable cost. This is because in many discussions that

concern the nature of autonomy the issue is *not* what psychological ca-
pacities a person must possess to have the *capacity* for autonomy, for it
is generally accepted that to be autonomous an agent must possess the
ability to engage in some form of second-order reflection of the sort that
Dworkin outlines. Instead, what is really of interest in discussions of au-
tonomy is the question of how the *exercise* of this psychological capacity
for reflection results in persons being autonomous with respect to their
desires and actions. Thus, in adopting this more global approach to au-
tonomy Dworkin is no longer offering an analysis of autonomy that is
congruent with the discussions in moral philosophy in which autonomy
plays a major role, for these discussions focus on the more localized ques-
tion of what makes a person autonomous with respect to her *particular*
desires or her *particular* actions.

Once Dworkin's more recent aims in developing an analysis of auton-
omy have been clarified, then, they can be seen to be distinct from the
primary aim of most autonomy theorists – namely, to provide an account
of what it is for a person to be autonomous with respect to her desires
and her actions. Yet this core aim of autonomy theorists is not satisfied
by Frankfurt's decision-based analysis of autonomy either, for it fails as a
successful response to three of the objections outlined above. First, the
mere fact that a person has *decisively* identified herself with a particular
first-order desire does not halt any possible problematic regress. This is
because, as Frankfurt later recognized, the Regress-cum-Incompleteness
Problem would still arise, given that one could still question whether the
person in question was autonomous with respect to this decision. Fur-
thermore, the Problem of Manipulation still poses difficulties for this
account because such a decision could still be the result of the agent's
succumbing to forces that are external to her. For example, she might
have been hypnotized into decisively identifying with a given desire.[28]
Finally, because a person can be manipulated into decisively identifying
herself with a particular first-order desire, the proponents of the *Ab Initio*
Problem/Problem of Authority can still question why such mental acts
are authoritative for her.

Frankfurt recognized that his analysis of autonomy was beset by these
three problems because it rested on the claim that a person became
autonomous with respect to her desires through endorsing them with
a "deliberate psychic event" – and one could always question whether
the person in question was autonomous with respect to this event. To
avoid these criticisms, Frankfurt developed a satisfaction-based analysis
of identification.[29] On this analysis, a person need not engage in any

"deliberate psychic event" for her to identify with her desires. Instead, on this analysis a person is autonomous with respect to a desire if he accepts it as his own – if he accepts it as indicating "something about himself."[30] In accepting a desire, a person will reflect on it to see if it is expressive of something about him. If it is, then he will form a higher-order attitude of acceptance toward it as part of himself. It is this *acceptance* of the desire that constitutes the person's *endorsement* of it, to use Frankfurt's "misleading" terminology from "The Faintest Passion."[31] The sense of endorsement that Frankfurt is using here, then, is the sense in which one might endorse the claim of an entity to be a member of a class, without thereby evaluating (either positively or negatively) the merits of the particular entity that is making the claim. Once a person has met the requirement that she reflectively endorse her first-order desires in this way, Frankfurt does not also require that she then reflectively endorse her attitude of endorsement, for, as he rightly notes, such a requirement would lead to a regress. Instead, Frankfurt holds that a person will identify with a first-order desire if she is *satisfied* with the higher-order attitude of endorsement (i.e., acceptance) that she has taken toward it. For Frankfurt, a person's being satisfied with his attitudinal set "does not require that [he] have any particular belief about it, or any particular feeling or attitude or intention. . . . There is nothing that he needs to think, or adopt, or to accept; it is not necessary for him to do anything at all." Instead, his being satisfied with his attitudinal set simply consists in his "having no interest in making changes" in it.[32] And this, notes Frankfurt, is important, for it explains why this analysis of identification as satisfaction is not subject to a problematic regress of the sort that beset his earlier analyses.[33] Here, then, a person will be autonomous with respect to his effective first-order desire if he is not moved to make changes in his motivational economy when he is moved to act by it, if he is *satisfied* with it.

Frankfurt's satisfaction-based analysis of autonomy is not subject to the Regress-cum-Incompleteness Problem for the reasons outlined above. Moreover, it is also not subject to the *Ab Initio* Problem/Problem of Authority. This is because Frankfurt has now clarified that a person's higher-order attitude of acceptance toward her lower-order desires does not possess any normative authority over them; instead, these attitudes are merely used by the person in question to assess whether her lower-order desires are to be regarded as being descriptively hers, whether they flow from her (broadly Lockean) self. However, this analysis of autonomy still faces the Problem of Manipulation. This is because a person could

unwittingly be hypnotized into possessing a certain first-order desire in such a way that he believes that it originates from within him. Given this belief, he would then both endorse this first-order desire and be satisfied with it, in Frankfurt's senses of these terms. This person would thus meet all of Frankfurt's most recent criteria for him to identify with his hypnotically induced desire – yet surely such a desire is one with respect to which its possessor is paradigmatically heteronomous.

II. NEW APPROACHES TO AUTONOMY

Christman's Historical Analysis

From the previous discussion, it might seem that the hierarchical approach to analyzing personal autonomy is doomed to failure, in large part because it appears inevitably to succumb to the Problem of Manipulation. Yet this assessment of hierarchical theories of autonomy needs to be qualified, for the focus of the past discussion was on Frankfurt's explicitly *ahistorical* approach to analyzing autonomy. Recognizing the difficulties that such an approach would have when faced by the Problem of Manipulation, Christman developed an explicitly historically based version of the hierarchical approach to analyzing autonomy. For Christman, an agent P is autonomous relative to some desire (value, etc.) at time t if and only if

  i. P did not resist the development of D (prior to t) when attending to this process of development, or P would not have resisted that development had P attended to the process;
  ii. The lack of resistance to the development of D (prior to t) did not take place (or would not have) under the influence of factors that inhibit self-reflection;
  iii. The self-reflection involved in condition i is (minimally) rational and involves no self-deception;[34]

and

  iv. The agent is minimally rational with respect to D at t (where minimal rationality demands that an agent experience no manifest conflicts of desires or beliefs that significantly affect the agent's behavior and that are not subsumed under some otherwise rational plan of action).[35]

Unfortunately, as it stands, Christman's historical analysis of autonomy fails to provide either necessary or sufficient conditions for a person to be autonomous with respect to her desires. To see that this account does not

provide necessary conditions for a person to be autonomous with respect to her desires, imagine a child at time t whose mother wished him to learn to play the piano and who beat him if he did not practice.[36] As time passes and the child grows more proficient at playing, he discovers (at time t1) that his mother's belief that piano playing suited him was right, and he comes to love playing – even though he still repudiates the means by which his mother brought him to this position. Thus, even though at t1 this person rejects the process by which he was brought to desire to play the piano, at t1 (and onward) he appears to be fully autonomous with respect to this desire.[37]

Furthermore, just as it is not necessary for a person to meet Christman's condition for autonomy for her to be autonomous with respect to her effective first-order desire, neither is it sufficient for this. To see this, one must note that Christman accepts that a person is autonomous with respect to a desire D even if she came to possess it under the influence of factors that inhibit self-reflection, provided that exposure to such factors was autonomously chosen.[38] Now consider the case of a man who wishes to join an order of monks who strictly follow the teachings of St. Ignatius of Loyola. It is a feature of the Ignatian tradition that its monks are required to subordinate their wills entirely to that of their abbots. No room at all should be left for the exercise of free choice or rational critical reflection, for these simply make the monk vulnerable to the temptation of Satan.[39] Knowing this, at time t this man decides to join the Ignatian order, thus autonomously choosing to subject himself to factors that inhibit self-reflection – namely, those that are required for him to subjugate his will to that of his abbot. If he is successful in his attempts to subjugate his will in this way, this man will (at time t1) only desire that which his abbot tells him to desire; he will, in effect, have reduced himself to the status of an automaton. However, he will still meet Christman's criteria for him to be held to be autonomous with respect to the desires that he has at time t1. This is because (since he had faith in his abbot) he would not have resisted the development of the desires he had at t1 had he attended to their generative process, the reflection-inhibiting factors that prevented him from reflecting on his desires were those that he autonomously chose, and he was minimally rational and not self-deceived at t1 also. However, because the *only* desires that he has are those that his abbot instructs him to have, this monk is a paradigm of heteronomy, rather than autonomy. And, because this is so, then even if a person's possession of his desires meets Christman's conditions, this does not suffice for him to be autonomous with respect to them.

Bratman's Approach

Christman's historical approach to analyzing autonomy was intended to be a development of the hierarchical approaches of Frankfurt and Dworkin. In a similar vein, Michael Bratman developed his reasons-based analysis of autonomy after he leveled what he took to be fatal objections to Frankfurt's satisfaction-based analysis.[40] At first approximation on Bratman's account, a person is autonomous with respect to a desire if she *decides* to treat it as being reason-giving (in the sense of being end-setting) in the relevant circumstances.[41] Bratman recognizes, however, that a person's decision to treat a desire as reason-giving is not sufficient for her to be autonomous with respect to it. This is because an unwilling drug addict might decide to give in to his craving and take drugs simply because it is becoming too painful for him to continue to resist his urges for them.[42] Here, the addict decides to treat his desire for drugs as being reason-giving in the relevant sense of being end-setting – and yet it seems that he is not autonomous with respect to it. To avoid this difficulty, Bratman argues that the key to understanding why the grudging addict is not autonomous with respect to his desire for drugs is that this desire is "incompatible with the agent's *other* standing decisions or policies concerning what to treat as reason-giving."[43] To be autonomous with respect to a desire, then, one must not only decide to treat it as being reason-giving but must also be *satisfied* with it. For Bratman, this satisfaction will consist in one not having "reached and retained a conflicting decision, intention or policy concerning the treatment of one's desires as reason-giving."[44]

As well as avoiding the Regress-cum-Incompleteness problem, Bratman's reasons-based analysis of autonomy also avoids the *Ab Initio* Problem/Problem of Authority. This is because he bases his account of what constitutes a person's standing decisions, intentions, and policies by reference to his broadly Lockean account of personal identity, on which an agent helps "ensure appropriate psychological continuities and connections [to retain her identity over time] by sticking with and executing [her] prior plans and policies, and by monitoring and regulating [her] motivational structures in favor, say, of [her] continued commitment to philosophy."[45] Because a person's standing decisions, intentions, and policies are constitutive of her self, they do indeed possess the authority to play the role in Bratman's analysis of assessing which of a person's first-order desires she is autonomous with respect to and which she is not. Furthermore, one need not ask whether the person is autonomous

with respect to her standing decisions, intentions, and policies. This is because, on Bratman's account of autonomy, these cross-temporal mental states (at least partially) constitute her self, and so (for the reasons that Noggle outlines in his paper "Autonomy and the Paradox of Self-Creation," this volume) the question of whether she is autonomous with respect to them does not arise.

Yet although Bratman's reasons-based approach to analyzing personal autonomy avoids two of the primary difficulties that beset its hierarchical predecessors, it still appears to be subject to the Problem of Manipulation. To see this, consider again a person who has been hypnotized into both having certain desires and accepting these desires as his own. Just as this person satisfied Frankfurt's criteria for him to be autonomous with respect to his hypnotically inculcated desires, so, too, does he satisfy Bratman's criteria for him to identify with them. This is because, owing to his hypnosis, this person treats these desires as being reason-giving in the sense of being end-setting, and they do not conflict with any of his standing "decisions or policies concerning what to treat as reason-giving," for he has not formed any views concerning the status of any hypnotically inculcated desires that he might have. Bratman, then, is also committed to the view that this person is autonomous with respect to his hypnotically induced desires – and this view is false. However, given Bratman's broadly Lockean account of personal identity that undergirds his account of autonomy, he might have an answer to this – that in such cases, the person's desires do not flow from her self in the appropriate way.[46] To develop this line of response, Bratman would have to strengthen his criterion that a person's decision to treat a desire as being reason-giving not conflict with her standing decisions, policies, and intentions to the claim that it must be in accord with them, and also add in a historical component to Bratman's view to block any revised versions of the Problem of Manipulation that might be developed against this strengthened version of his account.[47] But this is certainly a promising line of inquiry to take.

Ekstrom's Coherentist Analysis

It appears from this discussion that the Problem of Manipulation is an especially difficult one to avoid, although Bratman's analysis of autonomy might be modified to do so. There is, however, an alternative approach to analyzing autonomy that is immune to this objection and that deserves wider attention. This is the coherentist approach Laura Waddell Ekstrom has developed and that she elaborates upon in her contribution to this

volume, "Autonomy and Personal Integration." Ekstrom draws on the same insight that led Frankfurt and Bratman to develop their satisfaction- and reason-based analyses of autonomy: that a person is autonomous with respect to those conative states that move her to act if these flow from her self. Yet rather than analyzing what it is for a person to be autonomous with respect to her *desires*, Ekstrom is concerned with offering an account of what makes a person autonomous with respect to her *preferences*. For Ekstrom, a preference "is a very particular sort of desire: it is one (i) for a certain first-order desire to be effective in action, when or if one acts, and (ii) that is formed in the search for what is good."[48] Ekstrom's concept of a preference is thus like Frankfurt's concept of a second-order volition, except that Frankfurt allowed that a person might form a second-order volition for any reason at all, whereas for Ekstrom a person forms a pref- erence for a first-order desire because he finds a certain first-order desire to be good.

In developing her original coherentist analysis of autonomy in her paper "A Coherence Theory of Autonomy," Ekstrom distinguished be- tween a person's "self" and her "true or most central self." For Ekstrom, a person's "self" consists of her character together with the power for "fash- ioning and refashioning" that character, where a person S's character at time t is constituted by "the set of propositions that S accepts at t and the preferences of S at t."[49] A person's "true or most central self," however, consists of that subset of these acceptances and preferences that actually cohere. Ekstrom offers three reasons why such cohering preferences and acceptances are to be accepted as the elements of a person's core self. First, she notes that such elements are long-lasting; they are "guides for action that will likely remain, since they are well-supported by reasons." Second, the attitudes that constitute a person's core self are "fully de- fensible" against external challenges; they are those attitudes that one will fervently cling to through time. Third, those preferences that are elements of one's core self will be those that one is comfortable owning; they will be those that one will act on wholeheartedly. With this in place, Ekstrom argues that a person is autonomous with respect to her pref- erences (they "are *authorized* – or sanctioned as one's own") when "they cohere with [her] other preferences and acceptances" and thus can be recognized as members of her true self.[50] Thus, concludes Ekstrom, when a person acts on an authorized preference (i.e., one that coheres with her true self) she will act autonomously, not only because she will be able to give reasons for her action, but also because she will be acting in a way that is characteristic of her.

Ekstrom's coherentist analysis of autonomy was developed to avoid the standard problems that beset its hierarchical predecessors, and in this it appears to succeed. It is not faced with either the Regress-cum-Incompleteness Problem or the *Ab Initio* Problem/Problem of Authority. This analysis also appears to avoid the Problem of Manipulation, for Ekstrom requires that any preference that a person is autonomous with respect to be one that the person concerned can justify by appealing to his core preferences – and because these core preferences are, on Ekstrom's view, constitutive of the agent, any manipulation of them will result in a new agent and not in a loss of autonomy for their possessor.

Ekstrom avoids these three problems by basing her coherentist analysis of autonomy on the insight that if a person is to be autonomous with respect to a preference, that preference must originate from that person's self in a particular, objective way. Yet accepting this insight need not lead one to adopt a coherentist model of personal autonomy. Bratman, for example, draws on this insight to develop his reasons-based account of autonomy. Similarly, Robert Noggle also draws on it in his paper for this volume to show that neither the Regress-cum-Incompleteness Problem nor the *Ab Initio* Problem is as troubling as autonomy theorists (both coherentist and noncoherentist) take it to be – and he does so without committing himself to any particular approach to analyzing autonomy.

Of course, that a noncoherentist approach to analyzing autonomy might be able to avoid the Regress-cum-Incompleteness Problem and the *Ab Initio* Problem/Problem of Authority just as well as Ekstrom's coherentist analysis does, does not undermine the theoretical appeal of her approach. What might undermine its appeal, however, is the possibility that it fails to provide sufficient conditions for a person to be autonomous with respect to her preferences. To see this, consider again the case of the Ignatian monk who has subjected his will to that of his abbot. Because the preferences that this monk has through the operation of his abbot's will would (in the ideal situation) cohere (in Ekstrom's sense) with those that constitute this monk's "true or most central self," they will be "authorized" for him – and so when he acts on them, he would, on Ekstrom's account, act autonomously. But because this monk is a paradigm of heteronomy, rather than autonomy, Ekstrom's early coherentist analysis of autonomy fails to provide sufficient conditions for a person to be autonomous with respect to his actions.

However, it must be admitted that rather than providing a counterexample to Ekstrom's analysis, the example of the Ignatian monk might simply indicate that the relationship between the concepts of autonomy

and authenticity is still unclear. If the property of "autonomy" is under-stood to apply to a person with respect to her desires and actions if they meet some criterion in addition to the negative criterion that she is not alienated from them, then, although one is likely to accept that the monk acts *authentically* when he is subject to the will of his abbot, one will deny Ekstrom's claim that he acts *autonomously*. Alternatively, if the property of "autonomy" is understood more broadly, such that a person will be autonomous with respect to her desires and actions if she is not alienated from them, then one is likely to accept, with Ekstrom, that the Ignatian monk acts autonomously. It is unlikely that the debate over the proper ex-tensions of these two terms will be decided by etymology, for the concepts of both autonomy and authenticity require that in some way a person's de-sires or actions flow from her self. Instead, it would be better to settle this debate by asking whether any more precision could be brought to bear on discussions of autonomy by adopting either a broader or narrower construal of this term.[51]

III. AUTONOMY, FREEDOM, AND MORAL RESPONSIBILITY

Given the above litany of difficulties that face the various contemporary analyses of autonomy, one might worry that despite the considerable de-gree of attention the concept of autonomy has received in recent years, no real progress has been made toward developing a theoretically satis-fying account of its nature. But this worry is unfounded for two reasons. First, by developing criticisms to current analyses of autonomy and thus seeing where their weaknesses lie, one can establish what features a theo-retically satisfactory analysis of autonomy must possess. For example, the vulnerability of both ahistorical analyses and subjectively based analyses (i.e., those that rely on the subjective evaluation of the desires in ques-tion by their possessor to determine if she is autonomous with respect to them) to the Problem of Manipulation indicates that an acceptable anal-ysis of autonomy should incorporate an *objective, historical* condition for a person to be autonomous with respect to her desires. That is, to avoid the Problem of Manipulation, an analysis of autonomy must require that for a person to be autonomous with respect to a desire, she must have come to possess that desire as a result of some particular historical process – and this process must not be one that is based on the person herself adopting a particular attitude toward the origins of the desire that is in question. Second, the more recent analyses of autonomy all share certain features in common that might indicate that they are starting to converge